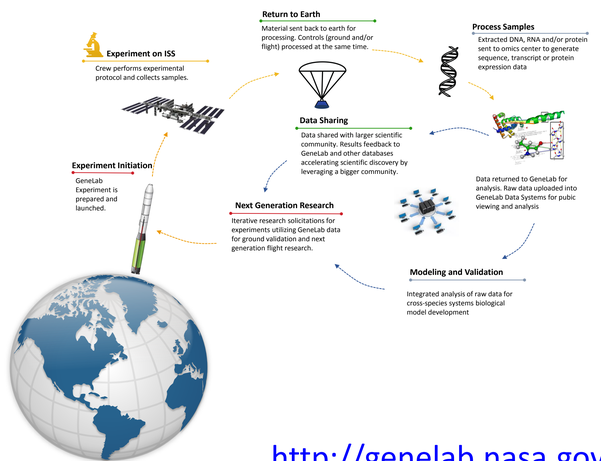
https://ntrs.nasa.gov/search.jsp?R=20150021762 2019-08-31T05:23:36+00:00Z

GeneLab: NASA's Open Access, Collaborative Platform for Systems Biology and Space Medicine

Daniel C. Berrios, MD MPH PhD[1], Terri G. Thompson, PhD[2], Homer W. Fogle, MS[2], Jon C. Rask, MS[2], Joseph C. Coughlan, PhD[2]
[1]University of California, Santa Cruz, CA; [2]NASA Ames Research Center, Moffett Field, CA

Purpose

NASA's mission includes furthering our understanding of biological systems through space-based research in order to improve life on earth and to enable the human exploration of space. To achieve these goals, NASA is investing in GeneLab, a multi-year effort to conduct biological and medical research in space, principally aboard the International Space Station (ISS). Through the GeneLab project, researchers will include high-throughput genomic, transcriptomic, proteomic or other "omics" assays as part of their experiments conducted on the ISS. The raw data from these assays will be stored in the GeneLab Data Systems (GLDS) currently being developed for the project. GeneLab intends to support "open science" research on the housed data sets, creating a multiplier effect on the science return from these experiments; the GLDS will serve all omics data without restrictions to the public. The system will ultimately include a biocomputation platform with collaborative science capabilities, to enable the discovery and validation of molecular networks that are influenced by space conditions.



http://genelab.nasa.gov/

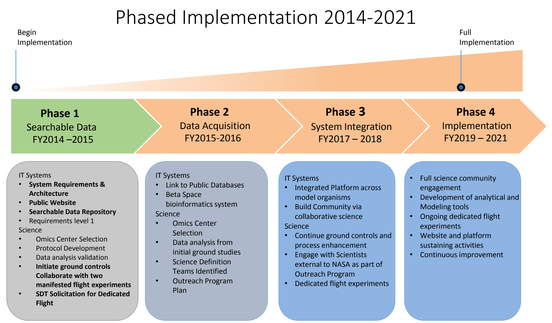
Design

NASA has chosen a phased capability implementation for the GLDS. The initial phase of the GLDS development effort emphasizes capabilities for the submission, curation, search, and retrieval of omics data sets. Important design considerations included:

- leveraging existing systems and systems components to deploy Phase 1 capabilities expediently;
- determination of optimal data set curation procedures including metadata representation and generation and quality control procedures); and
- balancing GLDS accessibility with user engagement and usage tracking capabilities.

Phase 2 will focus on interoperability and supporting federation of GLDS-housed data sets with externally-curated data in data set search, retrieval, and annotation functions.

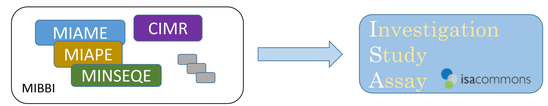
In Phase 3 we will develop a platform for computational biologists to execute and collaboratively develop analysis tools, integrate results of analyses with those of collaborating researchers, annotate data sets with interpretations, and share insights and hypotheses, building a knowledge base relevant to space biology and medicine.



Challenges

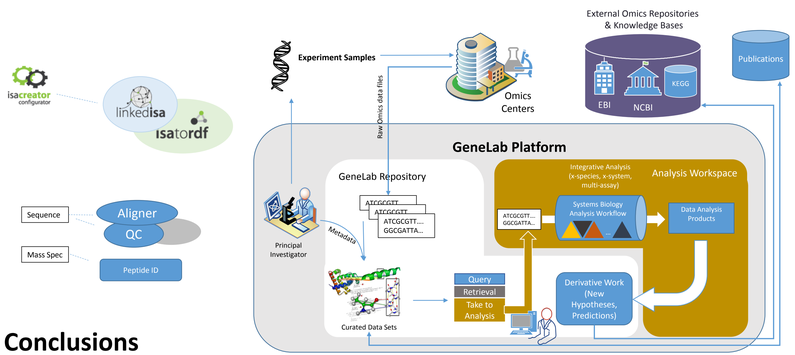
The GeneLab project and the GLDS face several challenges, including:

- Numerous and rapidly evolving standards for omics data, metadata, protocols and analyses



- Representing the unique aspects of space biology experiments as metadata and related data

- Coalescing access to a wide range of biocomputation tools, overcoming lack of interoperability
- Incentivizing engagement by the space biology/medicine research community



Conclusions

High-throughput systems for analyzing genomes and transcriptomes are generating vast amounts of biological data. Integrated analyses of these data with proteomic, metabolomic, physiologic, and phenotypic information holds the promise of rapid elucidation of complex molecular pathways, and of huge advancements in biological understanding and space medicine. Data Systems like the nascent GLDS are challenged with the representation, organization and integration of these highly complex data sets, and with providing researchers with the tools and environments they require for maximum utilization of the data in their analyses.